

# 領域間対応による両眼立体視

近藤秀樹 徐剛 辻三郎

大阪大学基礎工学部

現在、両眼立体視法における対応づけの研究の主流は、点ベース対応とエッジベース対応である。しかし、点ベースとエッジベースでの対応づけに比べ、領域間での対応づけの方が、対応づけの数が少なくてすむので一般に対応づけの効率が高い。対応づけの対象として領域が採用されないのは、領域分割の結果が必ずしも左右両画像間で一致せず、対応が取りにくいという問題があるためである。また、一つの誤対応が及ぼす影響が非常に大きいという問題もある。

本稿では、この領域分割の問題を解決するために、光源を制御しパターンを対象物体に投光する手法を提案している。パターンを投光することにより、簡単な二値化を行なうだけで領域を分割することが可能である。なお、このときの領域は実シーン中の領域ではなく、撮像された画像の明るい部分を指す。各領域は、様々な位置、特徴(面積、周囲長 e.t.c.)を持つことが対応づけにおいて望ましく、そのための投光パターンはランダムドット画像をガウスフィルタリングする事で作成できる。

対応づけは各領域の特徴を基に行なう。初めに、重心の座標を基に対応候補を探索し、次に各特徴を比較することで決定する。領域は、面積、周囲長など多くの情報を持つので対応づけは正確に行なわれ、また領域の数はエッジ点などに比べ非常に少ないので高速に行える。最後に、対応関係が決定された領域間で境界点を対応づけし視差情報を算出する。

## A Region-Based Stereo Algorithm

Hideki KONDO, Gang XU and Saburo TSUJI

Faculty of Engineering Science,  
Osaka University  
Toyonaka, Osaka, 560 Japan

This paper describes a stereo algorithm that works by projecting a high contrast random pattern onto the scene, thresholding intensities in the left and right images, and matching connected regions between the two images. Boundary points of the regions are then matched to yield disparities for all non-horizontal boundary points. One real scene example (coffee cups scene) is shown.

Although other stereo algorithm exist that rely on the projection of structured light patterns, this one is original in that it matches regions of random patterns. The algorithm is efficient.

## 1. はじめに

視覚は、人間が外界の情報を得る上で、もっとも重要な役割を持つ。その割合は約83%と言われ、他の感覚器官に対し、比類のないものとなっている。このような、すばらしい人間の視覚機能を機械で実現することが、コンピュータビジョン<sup>1)</sup>と呼ばれる研究分野の目標である。特に最近では、移動ロボットの視覚という観点から、三次元の世界を対象とする研究が、熱心に行われている。

いうまでもなく、画像は、三次元シーンが二次元平面上に写像され、縮退して記録されたものなので、一枚の画像からもとの三次元情報を導出することは、数学的には不可能である。それにもかかわらず、我々は、視覚により外界の様子を知り、その中で行動することができる。これは二つの眼を用いた両眼立体視、自分自身の動きや対象物の動きによる移動立体視、さらには、我々を取り巻く世界に関する知識を用いた推測などにより、総合的に三次元の情報を計測し復元しているものと考えられる。<sup>1,2,3)</sup>

一枚の画像から三次元情報を直接に導出することが数学的に不可能であるならば、人間の行っているように、同じシーンの異なる位置で取った二枚の画像を解析するのが一つの道であることを、だれでも容易に思いつくであろう。ステレオ視は、コンピュータビジョンの分野の最初からよく研究されてきた。現在、ステレオ視は、工業用ロボットや移動ロボットなど、いろいろな分野に広く応用されている。

現在のステレオ視の研究は特徴対応 (Feature-Based Matching)<sup>4)</sup> が主流である。すなわち、対応づけは左右両画像の抽出された特徴の間で行われる (例えば、エッジ点、エッジセグメント、領域など)。この種の手法は、撮像状況に影響されにくい、相関法に比べ、対応づけが一層安定、かつ正確である。対応づけを行う対象を、エッジ点、エッジセグメント、領域と代えることで、対応づけの個数を減らし、効率のよい、そして誤対応の少ない対応づけを行うことができる。しかし、それでも誤対応の可能性を無くすことはできず、また、対応づけの対象を高度なものにすればするほど、セグメンテーションに依存する割合が高くなり、かつ、一つの誤対応が持つ意味は大きくなる。対応づけの問題に関しては、Coarse-to-Fine 法<sup>23)</sup>、動的計画法、もう一台カメラを導入する三眼視などが提案されているが、いまだに決定的となる手法はない。

本稿では、ランダムなパターンを三次元物体に投光<sup>6)</sup> し、これを二台のカメラによって両眼立体視して三次元情報の計測を行う手法を提案している。従来の規則的な

パターン投光<sup>7)</sup> では、対応づけが困難であった。そこで、対応づけが容易に行えるパターンを作成する必要がある。これは、ランダムドット画像から容易に作成できる。次に、そのパターンを対象物体に投光することで、領域分割を容易にする。これは領域としてシーン中の領域ではなく、画像上で光の当たっている部分を採用することによる。対応づけは各領域の特徴を用いて行い、従来とは違った意味での両眼立体視法における領域間での対応づけを実現する。次に、対応関係にある領域の境界点を対応づけ、視差情報を算出する。

## 2. 領域間対応の原理

両眼立体視法では、対応づけを行う対象により、

- ポイントベース対応
- エッジベース対応
- 領域ベース対応

に大別される。

ポイントベースでの対応づけに比べ、エッジベースでの対応づけでは、対応候補の数は大きく減少し、またエッジセグメントの持つ特徴は、ポイントの場合に比べて多い。つまり、画像全体としては特徴は減少し、個々の特徴点 (この場合エッジセグメント) の持つ特徴量は増加する。その結果、対応づけは、より容易にかつ高速に行うことができる。

その延長上にあるのが、領域間対応の手法である。しかし、対応づけの対象として領域を考えると、領域分割という大きな壁に衝突する。もし、領域分割が完全に行われるのであれば、領域間対応に勝る手法はない。

本章では、この領域分割を容易にする手法について述べ、領域間での対応づけの原理を説明する。

### 2.1 領域分割

一般に、両画像間での領域分割の結果は一致しない。これは、左画像では一つの領域として存在するものが、右画像においては複数に分割されるなどといったことが起こるからである。実シーンをとらえた画像上では、正しい領域分割結果を得るためには、明るさ情報だけでなく、深さ情報も必要であるかも知れない。しかしまた、領域分割は深さ情報を算出するための前提ともなっている。これらの問題は、タマゴと鶏のどちらが先かという問題に似ており、解決することは非常に困難である。

領域分割の問題は、シーンに図2.1 に示すパターンを

対象物体に投光し、それを二台のカメラで撮像することで解決できる。強度の高い光を用いて投光し撮像すると、簡単な二値化を行うだけでパターン画像の明るい部分を容易に抽出することができる。注意すべきことは、ここでいう領域とは、シーン中の領域ではなくて、図2.1に示すパターンにより、光が投光されている部分のことである。

次に、8近傍を探索することで領域のラベリングを行う。このときに用いるラベリングは、領域にそれぞれ番号（明るさ）を与えることで実現する。

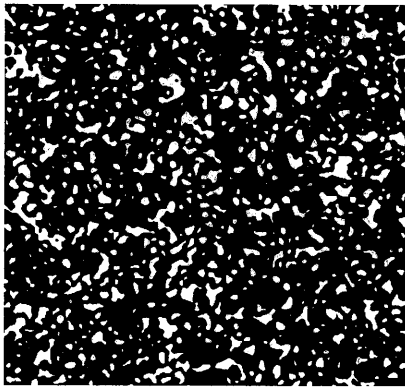


図2.1 パターン画像

## 2.2 領域の記述

領域分割により、閉じた領域が得られると、その形状、大きさなどについての大局的な特徴を求めることができる。本実験では、以下に示すような大局的特徴を用いて、各領域を記述する。

- 面積：領域内の画素数
- 周囲長：領域の境界点の画素数
- 重心の座標： $x = \Sigma x / \text{area}$   
： $y = \Sigma y / \text{area}$
- 外接長方形：領域内の  $x, y$  座標の最大値  
および最小値
- 領域の境界上の点の座標：チェーン符号

当然のことではあるが、各領域は上で述べた特徴のほかにも多くの情報を持つ。今回の実験では、上記の特徴を使ったが実際には必要のないものもある可能性があり、また別の特徴を使った方が効率的であるかも知れない。（可能性は小さいと考えられる。）この点は今後の課

題である。

## 2.3 探索範囲の限定

対応候補領域探索は、重心の座標をもとにして行う。ステレオ視での対応づけにおいては、エピーポラ拘束が大きなウェイトを占めるが、もしカメラと対象の間の距離がある程度正確に解っているとすると、エピーポラ線上の点について全てを探索する必要はなく、探索範囲をある程度限定することができる。

したがって、左画像に対して右画像で対応候補領域を探索するとすると、図2.2に示す実線の範囲（探索範囲(1)）を探索すればよい。点線の範囲（探索範囲(2)）についての詳しい説明は、次節に譲る。視線の上下に幅を持たせているのは、左右の重心は元々違うものであるため、また必ず存在する垂直視差に対処するため、そして量子化誤差によるためである。

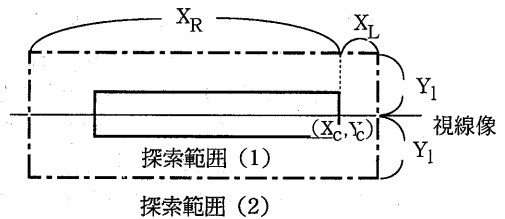


図2.2 対応候補領域探索範囲

$(X_c, Y_c)$ : 重心の座標

## 2.4 対応づけ

こうして対応候補領域が見つかったら、次に示す各条件を満足するかどうかを調べ、全てを同時に満たした領域を真の対応候補領域として決定する。

- $|\text{area}_L - \text{area}_R| / \text{area}_L \leq T_{\text{area}}$
- $|\text{perimeter}_L - \text{perimeter}_R| / \text{perimeter}_L \leq T_{\text{perimeter}}$
- $|(\max_{x_L} - \min_{x_L}) - (\max_{x_R} - \min_{x_R})| \leq T_x$
- $|(\max_{y_L} - \min_{y_L}) - (\max_{y_R} - \min_{y_R})| \leq T_y$

ここで、 $L, R$ はそれぞれ左画像、右画像を示す。また、各評価値  $T_{\text{area}}, T_{\text{perimeter}}, T_x, T_y$  は、視差の範囲、対象物体によって決定される。

実際の対応づけの処理においては、以下に示す対応策を用いることで、誤対応の問題、また複数対応の問題を

解決できる。

・対応づけの条件を厳しいものから、順次制限を緩めて複数回の対応づけを行う。(仮の対応づけ)

#### ・仮の対応づけ

深さの不連続点に沿っては、投影の関係上、左画像上では一つの領域が右画像上では複数に分割されたり、またその逆に、左画像上では複数の領域が右画像上では一つに見えるなどといったことが起こり得る。このとき、どちらの場合においても、重心の座標は大きく変化するはずである。視線像上だけを探索するのであれば、このような領域間の対応づけを行うことはできない。

そこで、左画像と右画像において、余り見え方の違わない領域間で、きびしい条件、狭い探索範囲のもとで先に対応づけを行ってやる。次に、探索範囲の上下左右に十分な幅を持たせてやること(図2.2:探索範囲(2))で、それらの領域間に仮の対応づけを行ってやる。この操作により、二つの画像間で大きく見え方の違う領域、あるいは左画像では一つの領域が、右画像では複数に分割されている領域などに、仮の対応候補を与えることができる。しかし、ここで注意すべきことは、これはあくまでも仮の対応づけであるということであり、対応づけは一対一で行われるので、上のような場合、当然ながら左画像における領域と、それに対応する右画像上での領域は完全に同じものとはいえない。(しかし、領域の一部は正しく対応づけられている。)

### 2.5 視差情報の獲得

二つの画像間で領域間の対応関係が求まると、境界上の点に対して視差を計算する。視差は、二つの画像間で対応関係にある点の間の座標値のずれであり、ほぼ距離

に反比例し、無限遠方で零となる。また、これらの視差はカメラモデルを定義するカメラ間隔およびレンズ焦点距離に比例する。

視差を算出するために、今度は境界上の点に対して、対応づけを行ってやる必要がある。両眼立体視法の性格上、水平方向のエッジに対しては、原理的に視差の算出は不可能なので(一般に両眼立体視では、カメラを水平に配置するためである。)、垂直方向のエッジに対してのみ、対応づけ、視差の算出を行う。

対応づけは、対応する領域間で視線像を引き、その視線像上の点に対して行なう。(図2.3)

### 3. 処理の手順と実験結果

本実験を行うにあたり、互いのエピポーラ線は、ほぼ水平線(カメラの走査線)と一致するようにカメラを配置した。つまり、光軸は互いに平行であり、かつ基線に垂直である。また、以下の実験においては、エピポーラ線は水平であると仮定している。

カメラの配置図を、図3.1に示す。

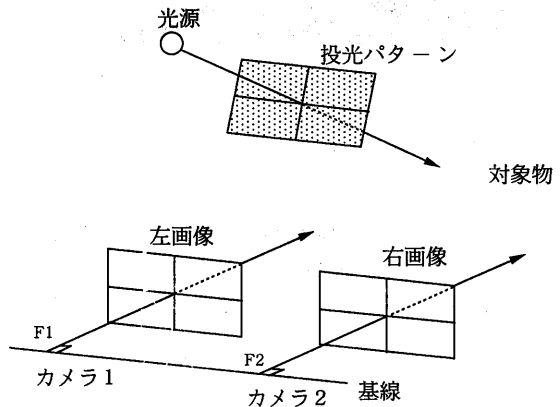


図3.1 カメラの配置

#### 3.1 パターン画像の作成

最初に、人間の視覚研究でよく用いられているランダムドット画像を用いることから出発する。これは、領域間での対応づけをするにあたり、様々な形状、大きさを持つ領域が、ランダムな位置に存在していれば、対応づけがより容易になると考えられるからである。

しかし、従来のランダムドット画像では、本実験で領域と考えている明るい部分(白いドットの集合)で、面積の小さいものが多すぎる。しかも一方では、面積の大

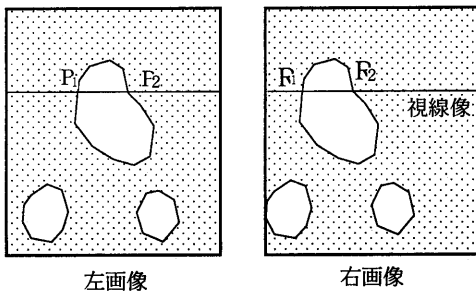


図2.3 境界上の点の対応づけ

きい領域も存在する。このことを周波数領域で言いかえると、それは各周波数成分が一様であるということである。

これは、対応づけをするにあたり大きな障害となり得る。小さな領域が多いということは、対応候補の数が増え、誤対応の確率が大きくなるということである。また、大きな領域が存在するという事は、シーン中の多数の領域、不連続な領域にまたがって投影される可能性があるということである。

そこで、これらの障害を克服するために、ガウスフィルタを用いて、周波数成分の平滑化（領域の大きさの平滑化）を行う。

この操作により作成したパターンを、図2.1に示す。

### 3.2 領域の分割と記述

パターンを投光する前の画像を図3.2に示し、以下に処理の手順を示す。

- [1] 2台のカメラにより撮像した2枚のステレオ画像に対して、2値化を行うことで、明るい部分と、暗い部分を明確に分割する。2値化を行なった結果を図3.3に示す。
- [2] 画像データをTospix-IIに転送し、ラベリングを行う。
- [3] 画面をラスタ走査することで、面積、外接長方形、重心の座標を計算する。
- [4] 8連結を用いて、周囲長を計算する。境界上の点の座標は、チェーン符号を用いて記述する。

なお、図3.3において左画像の領域は452個、右画像の領域は442個であった。

### 3.3 対応づけ

図3.3のステレオ画像に対して対応づけを行った。本実験においては、2回の対応づけを行い、以下の結果を得た。対応づけの回数、対象、およびパターンなどにより決定される。

先の例において、対応候補領域が一意に決定できなかった領域を、図3.4に示す。

図3.4を見ると明らかのように、左右両画像間で余り見え方の変わらない領域から対応づけが行われ、次に、見え方の大きく変化している領域間で、仮の対応づけが行われており、2.4での対策が実際に用いられているのが良く分かる。

また、2.4の対策を行ってもまだ、対応候補領域が複数存在する場合には、その中から、 $T_{area}$ ,  $T_{perimeter}$ の順に評価して、もっとも特徴が一致する領域を、仮の対応候補領域として決定している。

対応づけの結果、対応が決定できなかった領域は、一回目で86個、二回目で77個であった。

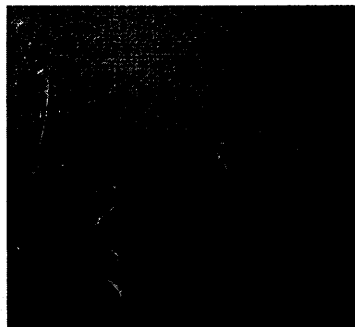
なお、本実験では、対応候補領域が一意に決定できなかった領域については、対応づけができなかったということにしている。そのため、左画像における領域で、右画像においては全く同じ二つの領域として、別々に存在する場合など、対応づけがとれない。しかし、このようなことは、非常に低い確率（深さの不連続な点に沿ってしか起こらず、しかも、面積、周囲長などが全く同じに分かれることなど皆無に等しい。）でしか起こらないので、問題とはならないであろう。

### 3.4 視差の算出

対応関係にある各領域間で、2.5に示したように境界上の点（垂直エッジに関してのみ）を対応づけすることで、

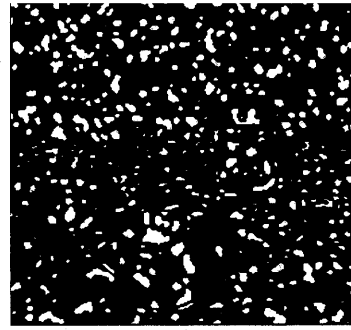
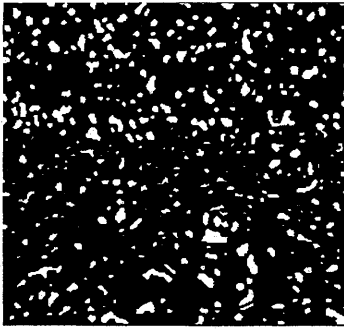


(a) 左画像



(b) 右画像

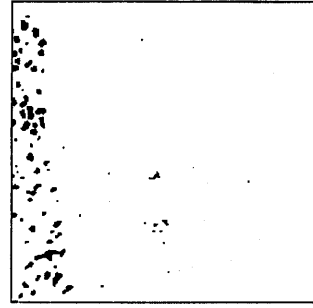
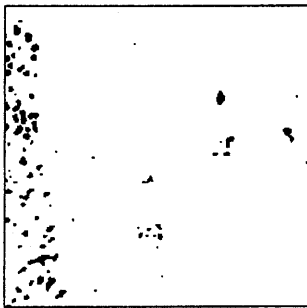
図3.2 ステレオ画像（パターン投光前）



(a) 左画像

(b) 右画像

図3.3 ステレオ画像（パターン投光後）



(a) 対応づけ1回目

(b) 対応づけ2回目

図3.4 対応候補が決定できない領域（左画像）

これらの境界点の間の座標のずれ（視差）を計算する。  
 このとき、計算された視差の値の中で、同じ領域の他の境界点における視差の値から、大きく値がかけ離れたものは、ノイズ（仮の対応づけによる影響）であるとして除去する。これは、図3.5に示すように、視差の値でヒストグラムを描き、図に示す範囲にあるものを除くことで実現する。除去対象範囲は、測定対象、測定精度に依存

して決定されるものであり、複雑なシーンでは、除去対象範囲は小さくなるが、その場合は多くのノイズを含むと考えられる。

つまり、この操作によって疑わしい情報は、全て除き、

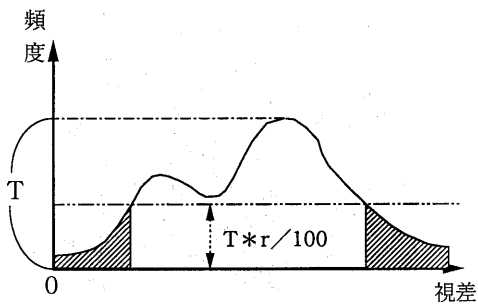


図3.5 視差のノイズ除去範囲

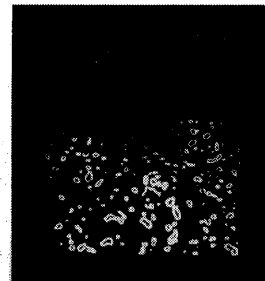


図3.6 視差図

正しいと思われる情報のみを採用することができる。

図3.6に、算出された視差を明るさで表した視差画像を示す。なお実験の結果、最大視差は55pixels、最小視差は29pixels、であった。

## 4. 検討と今後の課題

### 4.1 ランダムドット画像を用いることの正当性

図4.1に撮像された左右両画像における各領域の面積、周囲長の分布を示す。

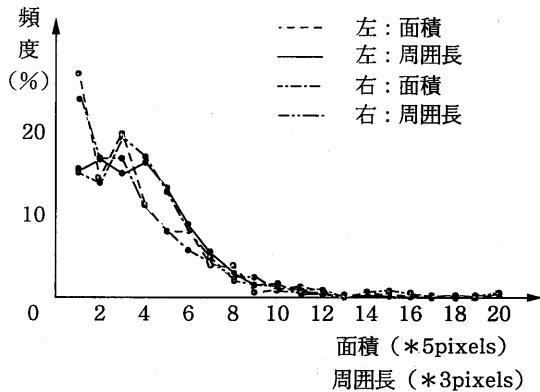


図4.1 面積、および周囲長の分布

図4.1に示す通り、面積及び、周囲長はともに広く分布をしている。つまりこれは、大きさ、形状などの様々な領域が、パターン画像中に存在していることを証明するものであり、対応づけを行うにあたり望ましい性質であると言える。なぜなら、様々な大きさ、形状を持つ領域が一樣に分布しているということは、誤った対応候補が少なくなるということを表しているからである。

また、図2.1を見ると明らかなように、パターン画像において、水平方向にはほとんど似た領域が存在しない。このことも、対応づけにおいて望ましい性質であると考えられる。

### 4.2 仮の対応づけの検証

図4.2に仮の対応づけの例を示す。

仮の対応づけが取られるのは、次の三通りの場合であると、考えられる。

- 1) パターン画像の投光、カメラの配置などの影響の

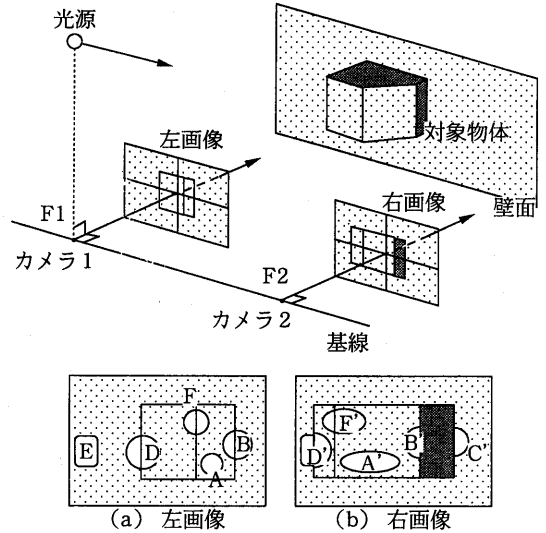


図4.2 仮の対応づけの例

ために、左右両画像間で大きく見え方が変化しているか、または、シーン中の連続した複数の領域に跨って投影された領域間の対応づけ。

(実は、正しい対応である。)

- 2) 左画像では一つの領域が、右画像においては、複数に分割されている場合、およびその逆の場合。(不連続なシーン中の領域に、一つの明るい部分が投影されることによる。)

- 3) 左画像では複数の領域が、右画像においては、一つに合体している場合、およびその逆の場合。(オクルージョンによる。)

図4.2において、領域Aと領域A'、および、領域Fと領域F'の関係は1)にあたり、領域Bと領域B'、領域Cと領域C'の関係は、2)にあたる。また、領域D、領域Eと領域D'の関係は、3)にあたる。

図3.5を見ると明らかなように、これらの領域間では、仮の対応づけが行われており、本実験における対応づけのアルゴリズムの正当性は、実証された。

### 4.3 測定対象と測定精度

測定精度については、パターン画像作成時に関係し、ガウスフィルタのマスクのサイズ(これは、 $\sigma$ の値に依存する。)によって決定される。

マスクのサイズが小さければ小さいほど、パターン画像中の領域の個数が増え、また、個々の領域の面積は、全体的に小さいものとなる。領域の数が増えることによ

り、正しい対応づけが完全に行われたと仮定すると、密な視差情報を得ることができ、複雑なシーンに適用することが可能となる。

しかし、一方では領域の数が増え、しかも個々の領域の大きさが小さくなり、似てくることで、誤対応の可能性が高まる。このため、測定精度と、密な視差情報とは、二律背反となる面が強い。

また、3.4で述べた視差情報のノイズ除去においても同じことがいえる。複雑なシーンに、パターン画像を投影した場合、当然ながら領域の各々の境界点の持つ視差は、様々な値を取ることが予想される。また、ここで仮の対応を取ることの影響が現れる。そのため、うまくノイズを除去することができなくなる。これは、各領域の面積を小さなものとすることで一部改善できるが、そのことにより、今度は誤対応が多くなってしまう。

本実験では単純なシーンを用いているため、ある程度大きな領域が散在するパターン画像で、十分な情報を得ることができた。しかし、機械部品の三次元形状測定などにおいては、このパターン画像では、不十分であると考えられる。

以上のことより、

- ・ガウスフィルタのマスクのサイズと、二値化のしきい値決定に何等かの評価関数を導入する。
- ・対応づけの拘束を厳しくする。
- ・一対多、および多対一の対応づけを行う。(4.3参照)
- ・ノイズの除去方法を改善する。

などの対策について、考慮する必要がある。

また、画像全体に対しての三次元情報が得られないなどの点を考慮する必要がある。

#### 4.3 一対多、および多対一の対応づけ

##### (部分的な対応づけ)

本実験を実現したアルゴリズムでは、仮の対応づけなどの考え方を導入し、一対一の対応づけを行っている。

しかし、パターンを投光し、それを撮像した画像では、4.2で述べた2) や、3) のような領域が存在する。

そのため、一対一の対応づけでは、(仮の) 対応関係にある領域間で三次元情報を求めても、その情報の全てが正しいと言うわけではなく、3.4に示すようにノイズの除去を行う必要がある。しかし、全てのノイズを除去できるといった保障はどこにもない。

また、左画像に対し右画像で対応候補を選択するとすると、4.2の(2)の場合、右画像において、複数に分割された各領域は、仮の対応づけができた一つの領域を除いたとしても、左画像の他の領域の誤った対応候補と成

り得る。

これは、投影パターンが既知のものであるので、知識の利用などにより、実現できるのではないかと考えられる。

## 5. おわりに

本論文では、従来と違った意味での領域間対応を実現する手法について述べた。ランダムドット画像を用いて作成したパターン画像を、対象物体に投光することで、領域間での対応づけの際に、大きな問題となっていた領域分割を解決した。

対象物体にランダムなパターンを投光し、撮像したステレオ画像間で、領域ベースでの対応づけの実験を行い、満足する結果を得た。

画像中のすべての点に対しては三次元情報が得られない、適用範囲が限定されるなどの問題点は残るが、実用化できるものと考えられる。

#### 参考文献

- 1) D. Ballard and C. Brown: "Computer Vision", Prentice Hall, Inc., 1982
- 2) W. Grimson: "A Computer Implementation of a Theory of Human Stereo Vision", Phil. Trans. Roy. Soc., London, Vol. B292, pp.217-253, 1981
- 3) D. Marr and R. Poggio: "A Computational Theory of Human Stereo Vision", Proc. Roy. Soc., London, Vol. B204, pp.301-338, 1979
- 4) W. Grimson: "Computational Experiments with a Feature Based Stereo Algorithm", IEEE Trans. PAMI, vol.7, No.1, pp.17-34
- 5) Medioni and R. Nevatia: "Segment-based Stereo Matching", Computer Vision, Graphics and Image Processing, Vol.31, No.1, pp.2-18, 1985
- 6) H. Nishihara: "PRISM: A Practical Realtime Imaging stereo Matcher", Proc. 3rd Conf. on Robot Vision and Sensory Controls, pp.121-129, 1983
- 7) K. Sato and S. Inokuchi: "Range Imaging System Utilizing Nematic Liquid Crystal Mask", Proc. 1st Int. Conf. Computer Vision, 1987, pp. 657-661